ANNALES
UNIVERSITATIS MARIAE CURIE-SKŁODOWSKA
LUBLIN — POLONIA
VOL. V, 3 SECTIO H 1971

Instytut Ekonomii Politycznej i Planowania Gospodarczego
Wydział Ekonomiczny UMCS

Sławomir KOZŁOWSKI

Pewne problemy związane ze stosowaniem metody różnic J. Czekanowskiego w badaniach przestrzennych

Некоторые проблемы, связанные с применением метода разностей Я. Чекановского в пространственных исследованиях

Einige mit der Anwendung der Differentialmethode von J. Czekanowski in Raumforschungen zusammenhängende Probleme

W ostatnich latach coraz częściej stosuje się do badań przestrzennych taksonomiczną metodę różnic J. Czekanowskiego[1] i metody pokrewne. Dotyczy to szczególnie badań ekonomiczno-rolniczych, choć również i inne dyscypliny zaczynają posługiwać się tą metodą.[2] Z zastosowaniem

---

[1] Chciałbym tu zwrócić uwagę, że — moim zdaniem — powinno się zróżnicować stosowaną terminologię zależnie od używanego w metodzie wzoru. Najczęściej metodę J. Czekanowskiego określa się mianem metody różnic przeciętnych. Stosowanie generalnie dla wszystkich wzorów przymiotnika „przeciętnych" jest niewłaściwe. Nazwę metoda różnic przeciętnych należy pozostawić dla wzoru, w którym występuje dzielenie przez liczbę cech — n, a więc występują rzeczywiście różnice przeciętne:

$$W_{1,2} = \sum_{j=1}^{n} \left| \frac{a_{1j} - a_{2j}}{n} \right|$$

Jeśli stosujemy wzór bez dzielenia przez liczbę cech, nie mamy do czynienia z różnicami przeciętnymi. Wówczas nie powinno się stosować terminu metoda różnic przeciętnych, lecz metoda różnic, czy ewentualnie różnic sumarycznych

$$R_{1,2} = \sum_{j=1}^{n} \left| a_{1j} - a_{2j} \right|$$

[2] Na przykład A. Fajferek (*Region ekonomiczny i metody analizy regionalnej*, PWE, Warszawa 1966), zastosował tę metodę do badań regionalnych, zaś Z. Gługiewicz (*Zastosowanie taksonomicznej metody różnic przeciętnych przy próbie wydzielenia rejonów klimatyczno-glebowych*, „Przegląd Statystyczny" 1961, nr 1) — do rejonizacji przyrodniczo-rolniczej.

tej metody do badań przestrzennych wiążą się — istotne moim zdaniem — problemy, które będę starał się przedstawić w poniższych uwagach.

Metoda różnic J. Czekanowskiego została opracowana dla badań antropologicznych. To samo dotyczy pokrewnych jej metod taksonomicznych, jak np. metody dendrytów, kwadratów różnic (Henzla) czy podobieństw. Przeniesienie jej do — odrębnych bądź co bądź — dyscyplin ekonomicznych pociąga za sobą określone skutki. W badaniach antropologicznych, jak również innych badaniach biologicznych (np. zootechnicznych)[3], przy pomocy tych metod wyszukujemy — ze zbiorowości osobników danego gatunku czy też rasy — grupy osobników jak najmniej różniących się między sobą pod względem wybranych cech diagnostycznych. Wszystkie osobniki badanej zbiorowości traktowane są tu jednakowo, co nie budzi zastrzeżeń. Mimo pewnych różnic należą one do danego gatunku czy rasy. Cechy typologiczne (diagnostyczne) są — w większości przypadków — cechami mającymi bezwzględny charakter, jak np.: wysokość, ciężar, wysokość i szerokość czaszki itp. Występujące obok tego wielkości relatywne (np. stosunek szerokości do wysokości czaszki) nie wprowadzają zasadniczych zmian, są one bowiem stosunkowo wyrównane.

Podsumowując powyższe można stwierdzić, że zbiorowości osobników w tych badaniach są stosunkowo wyrównane, a jednakowe traktowanie wszystkich osobników jest logicznie uzasadnione.

Pod tym względem badania przestrzenne, różnią się od wyżej omówionych dość zasadniczo. W badaniach tego typu miejsce poszczególnych osobników zajmują z reguły jednostki podziału administracyjnego danego terytorium. Zależnie od celu i zakresu badań mogą to być gromady, powiaty lub województwa. Opieranie się na takich jednostkach jest koniecznością, gdyż jedynie dla nich możemy otrzymać niezbędne do badań dane statystyczne. Dla każdego jest jednak jasne, że zbiorowość składająca się z pewnej liczby jednostek administracyjnego podziału jest bez porównania mniej wyrównana, niż analogiczna zbiorowość osobników biologicznych. Dotyczy to w pierwszym rzędzie wielkości powierzchni tych jednostek oraz liczby ludności. Wielkości te wahają się w znacznych granicach (tak np. powierzchnia użytków rolnych powiatów województwa lubelskiego waha się od 41,1 do 148,0 tys. ha; włączając zaś do analizy miasta wyłączone z powiatów rozpiętość ta jeszcze wzrośnie — od 2,9 do 148,0 tys.

---

[3] Por. prace: T. H e n z e l: *Metoda różnic i metoda kwadratów różnic*, „Przegląd Antropologiczny" 1953, t. 19; J. P e r k a l: *Taksonomia wrocławska*, „Przegląd Antropologiczny" 1953, t. 19; S. S c h m i d t i inni: *Zastosowanie metod taksonomicznych w zootechnice*, „Zeszyty Naukowe WSR w Krakowie" 1963, seria Zootechnika, z. 4.

ha).[4] Tak wielkie różnice w zbiorowości osobników biologicznych są niezmiernie rzadkie. To powoduje, że — w przeciwieństwie do badań biologicznych — w badaniach przestrzennych stosowanie miar (wielkości) bezwzględnych staje się niecelowe. Tego typu wielkości są bowiem całkowicie nieporównywalne. Stąd w badaniach przestrzennych cechy diagnostyczne są podawane w postaci wielkości relatywnych, w badaniach ekonomiczno-rolniczych najczęściej w przeliczeniu na 100 ha użytków rolnych, w regionalnych — na powierzchnię geodezyjną, na 1000 mieszkańców itp.

Fakt stosowania takich relatywnych wielkości w taksonomicznej metodzie różnic wymaga ustosunkowania się do pewnych zagadnień. Celem metody — w badaniach przestrzennych — jest połączenie w grupy (regiony, rejony, strefy) tych obok siebie położonych jednostek, które pod względem wybranych cech różnią się możliwie najmniej. A więc przykładowo w rejonizacji ekonomiczno-rolniczej łączymy jednostki o możliwie najbardziej zbliżonej do siebie strukturze użytków rolnych, upraw, obsadzie zwierząt — wszystko w przeliczeniu na 100 ha użytków rolnych.

Właściwie dobrane cechy typologiczne muszą się między innymi charakteryzować dużą zmiennością przestrzenną, bowiem jedynie wtedy istnieje możliwość podziału badanego obszaru. Dlatego z szeregu cech typologicznych wybieramy takie, dla których współczynniki zmienności ($V = \frac{\sigma}{M}$) są największe. Wielkości średniej arytmetycznej oraz odchylenia standardowego konieczne są również dla weryfikacji wyodrębnionych rejonów, przeprowadzanej przez większość autorów w oparciu o kształtowanie się współczynników zmienności. Obliczenie średniej niezbędne jest również dla przeprowadzenia standaryzacji (normalizacji)[5] różnic między wartościami cech, zgodnie z najczęściej przyjmowanym wzorem

$$R_{1,2} = \sum_{j=1}^{n} \left| \frac{a_{1j} - a_{2j}}{M_j} \right| \cdot 100$$

(dla przeprowadzenia normalizacji „zerojedynkowej” konieczne jest również obliczanie odchylenia standardowego). Pojawia się wobec tego problem, w jaki sposób należy obliczać wielkości odchylenia standardowego (σ) i średniej arytmetycznej (M).

Duża nierównomierność poszczególnych jednostek administracyjnego podziału sugeruje konieczność liczenia średniej oraz odchylenia jako wielkości ważonych. Praktycznie w żadnej ze znanych mi prac traktują-

[4] Por. *Rocznik statystyczny województwa lubelskiego 1969*, s. 122.

[5] Różni autorzy stosują tu różną terminologię używając określeń standaryzacja lub normalizacja.

cych o metodzie J. Czekanowskiego w badaniach przestrzennych nie spotkałem się z ustosunkowaniem do tego problemu. Z drugiej strony w pracach drukowanych nie przytacza się z reguły szczegółowych danych statystycznych, co nie pozwala na dojście w jaki sposób autorzy obliczali te wielkości.

Stąd wydaje się, że warto się zastanowić nad problemem, jak mają być obliczane wielkości średniej arytmetycznej i odchylenia standardowego w taksonomicznych badaniach przestrzennych. Za właściwy punkt wyjścia uważam tu cel metody, który sam — wydaje się — narzuca określony sposób postępowania. Celem metody jest — co warto jeszcze raz podkreślić — p o ł ą c z e n i e j e d n o s t e k z n i k o m r ó ż n i ą c y c h s i ę r e l a t y w n y m i w i e l k o ś c i a m i w y b r a n y c h c e c h niezależnie od tego czy jednostki te różnią się od siebie znacznie wielkością (do której odnosimy cechy), czy też nie. Nie interesuje nas zatem jaka jest r z e c z y w i s t a średnia danej cechy dla całego badanego obszaru i jaka jest r z e c z y w i s t a dyspersja tej cechy. Interesuje nas natomiast jak oscylują poszczególne r e l a t y w n e wielkości danej cechy wokół swojej przeciętnej obliczonej po prostu jako suma względnych wartości danej cechy dla wszystkich jednostek, dzielona przez liczbę tych jednostek $\left(\frac{\Sigma X}{N}\right)$.[6] Postępowanie powinno tu być więc identyczne jak przy samym obliczaniu różnic, gdy odejmując od siebie dwie wielkości względne $(|a_{1j} - a_{2j}|)$ abstrahujemy całkowicie od bezwzględnych ich wartości. Wielkości średniej i odchylenia obliczone w ten sposób są niejako wielkościami „fikcyjnymi", służącymi jedynie samej metodzie. Powinno to znaleźć swoje odbicie zarówno w stosowanej symbolice (symbole nie M, lecz np. M_{fikc}, σ_{fikc} czy też M_{metod}), jak i w terminologii. W miejsce stosowanego czasem określenia — średnia dla całego badanego obszaru (np.: średnio powiat, województwo — średnia), powinno się stosować — średnia „fikcyjna" względnych wartości cech dla wszystkich jednostek (gromad, powiatów).

W końcowej fazie badań, po zweryfikowaniu przeprowadzonej rejonizacji, kiedy przystępujemy do charakterystyki wydzielonych obszarów, powinno się jednak przechodzić do rzeczywistej wielkości średniej arytmetycznej. W takiej charakterystyce posługiwanie się średnimi „fikcyjnymi" jest niedopuszczalne. Wielkości te nie charakteryzują nam bowiem wydzielonych obszarów. Po dokonaniu podziału konieczne jest

[6] Z danych, jakie są zawarte w pracy A. F a j f e r k a (*op. cit.*, tab. 2, s. 62—63) można po przeliczeniu ich dojść do wniosku, że autor posługuje się tą właśnie metodą.

zatem przechodzenie na powrót do wartości bezwzględnych, by na ich podstawie móc scharakteryzować wyodrębnione obszary.

Należy bowiem jeszcze raz zaznaczyć, że średnie — określone tu jako „fikcyjne" — są konieczne przy posługiwaniu się samą metodą, a więc w trakcie podziału, nie mogą jednak charakteryzować wydzielonych obszarów.

Można się tu powołać też na stanowisko Z. Gługiewicza, który stosując standaryzację (normalizację) na średnią arytmetyczną tak uzasadnia dzielenie przez średnią:

„Średnia ma dla tych celów najwięcej zalet. Na pierwszy rzut oka można by wyrazić obawę, że przy wymaganym dużym obszarze zmienności cech (warunek 2) ma się do czynienia z niejednolitą zbiorowością, dla której tym samym nie można sensownie stosować średniej wartości cechy. Średnia nie jest tu jednak przesłanką opisu zbiorowości i nie jest podstawą do wyciągania jakichkolwiek wniosków w zakresie rejonizacji."[7]

Widoczne jest, że Z. Gługiewicz również posługuje się średnią jedynie dla wykorzystania jej w metodzie, ma więc ona charakter — jak już wspomniałem — wyłącznie metodologiczny.

Proces dokonania podziału, tj. uporządkowania diagramu jest w metodzie Czekanowskiego dosyć pracochłonny, gdyż dokonuje się go „[...] drogą kolejnych prób aż do opracowania zwartych, a zarazem odcinających się wyraźnie od siebie grup jednostek"[8]. Stąd niektórzy badacze poszukują metod usprawniających porządkowanie. I tak J. Liczkowski zwraca uwagę, że „pracę tę można usprawnić, wyznaczając kolejność wielkości decydujących cech taksonomicznych w drodze autopsji"[9]. Wydaje się, że w badaniach przestrzennych (rejonizacyjnych) nader pomocne może być wykorzystanie innego bardzo prostego kryterium wstępnego porządkowania. Polega ono na układaniu jednostek na diagramie w kolejności zgodnej z ich przestrzennym usytuowaniem na mapie. Kierunek, w którym się posuwamy (przenosząc poszczególne jednostki z mapy na diagram) jest przy tym zależny od kształtu badanego obszaru. Najkorzystniejszy jest kierunek zgodny z długością, tj. dłuższym wymiarem badanego terytorium. Wykorzystujemy przy tym grupowaniu zasadę, że rejony (obszary wydzielone) powinny być jednostkami przestrzennie zwartymi (oczywiście poza logicznie uzasadnionymi wyjątkami). Takie podporządkowanie nie jest oczywiście ostateczne, tym niemniej jest ogromnie pomocne i późniejsze zmiany ograniczają się naj-

---

[7] Por. Gługiewicz: *op. cit.*, s. 78.

[8] J. Liczkowski: *Badanie intensywności rolnictwa w ujęciu przestrzennym*, PWRiL, Warszawa 1964, s. 42.

[9] *Ibid.*, s. 43.

częściej do niewielu przesunięć. Tę metodę wykorzystałem z powodzeniem przy rejonizacji powiatu lubelskiego.

Dokonanie podziału badanego terytorium nie kończy postępowania badawczego. Z reguły większość badaczy przeprowadza następnie weryfikację dokonanego podziału dla stwierdzenia prawidłowości wyodrębnienia rejonów.

Co do sposobów, jakimi taką weryfikację się przeprowadza, to nie ma zasadniczo przystosowanych do metody różnic (przeciętnych), gdzie posługujemy się sumami różnic (lub różnicami przeciętnymi) szeregu wybranych cech diagnostycznych.

Część badaczy dla wykazania istotności różnic posługuje się testem istotności Studenta-Gosseta[10], tym niemniej stosowalność jego jest ograniczona z dwu względów. Po pierwsze jest to metoda nader pracochłonna, po drugie zaś daje jedynie możliwość porównywania ze sobą pary rejonów i to pod względem jednej cechy. Liczba obliczeń zwiększa się więc znacznie wraz ze wzrostem liczby rejonów i cech, a ponadto czytelność (jasność) uzyskanych wyników jest wątpliwa. Dwa rejony mogą się bowiem różnić istotnie pod względem trzech cech, nie różnić pod względem innych trzech i kompleksowy osąd jest — w tym przypadku — utrudniony.

J. Steczkowski[11] natomiast proponuje stosowanie nieparametrycznego testu chi-kwadrat (χ^2), który jednak również ze względu na dużą pracochłonność nie przyjął się powszechnie.

Z tego tytułu najczęściej stosowanym obecnie sposobem weryfikacji jest obliczanie współczynników zmienności (V) dla wyodrębnionych jednostek (rejonów, regionów) i porównywanie ich z takimi samymi wielkościami dla całego badanego obszaru. Przy prawidłowej rejonizacji (podziale) współczynniki dla rejonów winny być niższe niż dla całego delimitowanego obszaru. Taka sama metoda zalecana jest dla dołączania do wyodrębnionych rejonów jednostek terytorialnych ciążących do kilku (najczęściej dwu) rejonów. Prawidłowe rozwiązanie to takie, przy którym „[...] przesunięcie jakiegokolwiek powiatu [jednostki, na której opieramy rejonizację — uwaga moja S.K.] z jednego rejonu do drugiego nie spowoduje: 1) jednoczesnego zmniejszenia średnich współczynników zmienności ($\overline{V}_x$) dla obu rejonów, między którymi zachodzi przesunięcie powiatu; 2) większego spadku współczynnika zmienności ($\overline{V}_x$)

---

[10] Test ten stosuje W. Kwiecień w pracy: *Metoda modelowa w badaniach ekonomiczno-rolniczych*, PWRiL, Warszawa 1968, s. 53—55.

[11] Por. J. Steczkowski: *Zasady i metody rejonizacji produkcji rolniczej*, PWRiL, Warszawa 1966.

w jednym rejonie aniżeli wynosi wzrost tegoż współczynnika w drugim rejonie."[12]

J. Steczkowski opierając się na tym samym współczynniku zmienności wprowadza inny wzór:

$$D_j = \sum_{j=1}^{n} \frac{V_{oj}}{M_g}$$

gdzie V_{oj} — to zmienność całego delimitowanego obszaru pod względem j-tej cechy, zaś Mg — to średnia geometryczna zmienności tej cechy dla p określonych rejonów. Im wskaźnik (suma) jest wyższy, tym podział jest lepszy.[13]

Metody oparte na współczynniku zmienności, jakkolwiek mniej pracochłonne, zajmują — szczególnie w fazie porządkowania diagramu (często wielokrotnego) — sporo czasu. Wydaje się, że nie rezygnując z liczenia współczynników zmienności dla ostatecznej weryfikacji, można dla sprawdzenia poprawności podziału stosować znacznie prostszą metodę.

Punktem wyjścia dla tej uproszczonej metody weryfikacji jest stwierdzenie, że rejony to obszary jak najmniej różniące się wewnątrz (między jednostkami wchodzącymi w skład rejonu), a jak najwięcej różniące się na zewnątrz. Na diagramie więc rejony powinny zawierać jak najwięcej powiązań bliższych rzędów (np. przy skali czterostopniowej — pierwszego i drugiego), zaś poza swymi granicami zawierać tych powiązań jak najmniej. Odwrotnie powinny kształtować się proporcje co do powiązań dalszych rzędów (np.: trzeciego i czwartego), których wewnątrz rejonu powinno być możliwie jak najmniej, zaś na zewnątrz jak najwięcej.

Tę zasadę można wykorzystać dla utworzenia mierników poprawności rejonizacji (podziału). Takim miernikiem może być liczba powiązań I, II,..., n-tego rzędu wyrażona w procentach w stosunku do wszystkich możliwych powiązań wewnątrz rejonu (wyodrębnionej grupy), oraz takie same procentowe powiązania na zewnątrz rejonu (tj. między jednostkami wchodzącymi w skład rejonu a pozostałymi jednostkami). Dla łatwiejszej czytelności można je przedstawić w postaci następującej tabeli (tu przyjmuję skalę czterostopniową, gdzie pierwsze dwa stopnie to powiązania silne, drugie dwa — powiązania słabe) — tab. 1.

Na podstawie wielkości, jakie otrzymamy w ten sposób, możemy sądzić o prawidłowości dokonanego podziału.

Oczywiście od sporządzającego diagram, od przyjętej skali i oceny powiązań, zależeć będzie, którego rzędu powiązania uznamy za silne,

[12] Gługiewicz: *op. cit.*, s. 82.
[13] Steczkowski: *op. cit.*, s. 95.

Tab. 1

	Powiązania w % ogółu możliwych	Wewnątrz rejonu	Poza rejonem
Powiązania bliskie	I rzędu II rzędu		
	I i II rzędu		
Powiązania dalsze	III rzędu IV rzędu		
	III i IV rzędu		

a więc te, które powinny się skupiać w wybranych rejonach, którego zaś za luźne, a więc występujące poza rejonami. Tu przykładowo podaję skalę czterostopniową z dwoma powiązaniami silnymi i dwoma luźnymi; opieram to na skali, jaką zastosowałem dla rejonizacji produkcyjno-rolniczej powiatu lubelskiego. U innych autorów kwestia zaliczenia powiązań do silnych i luźnych — ze względu na określoną przyjętą skalę — może wyglądać inaczej. I tak w dwóch wersjach rejonizacji w pracy W. Kwiecienia[14] — mimo czterostopniowej skali — jedynie powiązania I rzędu można uznać za silne, pozostałe za słabe.

Wielkości z tabeli 1 można przedstawić w formie jeszcze bardziej syntetycznej, w łatwo czytelnej postaci jednoliczbowego wskaźnika. Taki syntetyczny, porównywalny wskaźnik — można określić go jako wskaźnik zagęszczenia (koncentracji) powiązań — może mieć, przy przyjętej tu czterostopniowej skali, następującą postać:

$$W_{zp} = \frac{P_{1w} + P_{2w}}{P_{1z} + P_{2z}} + \frac{P_{3z} + P_{4z}}{P_{3w} + P_{4w}}$$

gdzie W_{zp} — wskaźnik zagęszczenia (koncentracji) powiązań; p_{1w}, p_{2w}, p_{3w}, p_{4w} — powiązania I, II, III, IV rzędu występujące wewnątrz grupy (rejonu) i wyrażone w odsetkach wszystkich możliwych powiązań wewnętrznych; zaś p_{1z}, p_{2z}, p_{3z}, p_{4z} — powiązania I, II, III i IV rzędu występujące na zewnątrz rejonu (tj. między jednostkami wchodzącymi w skład rejonu a pozostałymi) i wyrażone w procentach wszystkich możliwych powiązań zewnętrznych.

Najogólniej (abstrahując od skali) wskaźnik taki ma postać:

$$W_{zp} = \frac{\text{wewn. pow. bliż. rzędów}}{\text{zewn. pow. bliż. rzędów}} + \frac{\text{zewn. pow. dalsz. rzędów}}{\text{wewn. pow. dalsz. rzędów}}$$

gdzie oczywiście wszystkie powiązania wyrażone są w procentach wszelkich możliwych powiązań.

[14] Kwiecień: *op. cit.*

Wskaźnik taki w postaci jednej liczby charakteryzuje nam „siłę”, z jaką wydziela się dany rejon z otaczającego obszaru (badanego). Im wyższy wskaźnik tym lepiej wydzielony rejon. Przesuwając poszczególne jednostki z rejonu do rejonu możemy badać jak wpływa to na kształtowanie się tegoż wskaźnika zagęszczenia powiązań w rejonach i wybrać optymalny wariant, przy którym wskaźniki przybierają najwyższe wartości.

Wskaźnik ten ma więc szczególne znaczenie dla wyboru optymalnego podziału obszaru przy ustalonej już skali. Tym niemniej można nim operować również przy ustalaniu skali, przy przesuwaniu granic klas szczególnie w niezbyt wielkim zakresie i niezmienionej liczbie klas.

W diagramie Czekanowskiego — ze względu na jego konstrukcję — na przekątnej układają się zerowe różnice danych jednostek samych z sobą. Wydaje się celowe przy stosowaniu wskaźnika zagęszczenia powiązań przedstawionego powyżej eliminowanie tychże powiązań (jednostek samych z sobą) z ogólnej liczby powiązań wewnętrznych. Oczywiście liczba powiązań pierwszego rzędu ulega także zmniejszeniu o powiązania ułożone na przekątnej. I tak jeśli przykładowo rejon składa się z 10 jednostek, to nie odnosi się powiązań wewnętrznych różnych rzędów do wszystkich powiązań w liczbie 100, lecz eliminuje 10 powiązań I rzędu ułożonych na przekątnej (jednostek samych z sobą) i odnosi wszystkie powiązania do 90. Takie eliminowanie przeprowadzamy z dwu zasadniczych powodów: 1— powiązania jednostek samych z sobą nie charakteryzują żadnych cech rejonu, występują one na przekątnej zawsze, co wynika z konstrukcji diagramu i będą wobec tego występowały również przy braku jakichkolwiek połączeń bliższych rzędów między jednostkami; 2— mają one różny wpływ na strukturę powiązań w rejonie, stanowią przykładowo w rejonie składającym się z 10 jednostek tylko 10%, przy rejonie z 4 jednostek już 25% wszystkich wewnętrznych powiązań. Nieeliminowanie ich zatem wpływałby korzystnie na kształtowanie się wielkości wskaźnika zagęszczenia w rejonach małych, niekorzystnie dużych. Przy takiej samej strukturze powiązań zewnętrznych i wewnętrznych, poza powiązaniami w rejonie ułożonymi na przekątnej, w rejonie małym (składającym się z mniejszej liczby jednostek) wyższa będzie pierwsza część wskaźnika zagęszczenia powiązań (pierwszy składnik sumy mówiący o koncentracji powiązań bliższych rzędów w rejonie), a zatem wyższy będzie również cały wskaźnik W_{zp}. Wobec tego wskaźnik obliczany z wyeliminowaniem tych powiązań jest bardziej prawidłowy.

Wielkość wskaźnika liczona zgodnie z podanym tu wzorem nie może być obliczona jedynie w tym przypadku, jeśli na zewnątrz rejonów nie istnieją powiązania bliższych rzędów bądź też wewnątrz nie istnieją

powiązania dalszych rzędów (np. przy przyjętej tu skali — III i IV) lub oba te przypadki występują jednocześnie ($W_{zp} = \infty$). Należy jednak zaznaczyć, że takie przypadki są w rejonizacji opartej na kilku cechach bardzo rzadkie i dowodzą wtedy — o ile tylko granice, co mamy możność kontrolować na diagramie, wyznaczone są prawidłowo — wręcz idealnego wydzielenia rejonów. Zawsze bowiem trzeba kierować się logiką w wyodrębnianiu rejonów. Można tak rozszerzyć granice rejonu na diagramie, żeby zmieścić w nim wszystkie bliższe powiązania — nawet jeśli są one pojedyncze — na zewnątrz pozostawiając tylko powiązania dalszych rzędów i w ten sposób doprowadzić do ukształtowania się wskaźnika równego nieskończoności. Takie postępowanie nie ma jednak sensu. Kontrolą prawidłowości postępowania jest tu w pewnym stopniu struktura wskaźnika koncentracji powiązań. Nie powinno być tak, że np. część pierwsza (pierwszy składnik sumy) jest mniejsza od jedności, a jednocześnie druga równa się nieskończoności. Zarówno pierwsza część wskaźnika — świadcząca o zagęszczeniu powiązań bliższych rzędów w rejonie, jak i druga — świadcząca o koncentracji dalszych powiązań poza rejonem, powinny być wielkościami zbliżonymi do siebie.

Zaznaczyć też warto, że kształtowanie się $W_{zp} = \infty$ może wynikać z nieprawidłowego wyboru skali (np. bardzo ostre granice pierwszych stopni podobieństw, co może powodować wydzielenie dużej liczby niewielkich rejonów o braku powiązań zewnętrznych lub też przyjęcia bardzo luźnych granic dla pierwszych stopni, co może powodować, że cały badany obszar staje się jednym dużym rejonem).

Obok łatwości obliczeń na dobro wskazanej tu metody należy zapisać także to, że dla przesuwania jednostek między rejonami w celu jak najlepszego podziału w oparciu o ten wskaźnik, nie trzeba sporządzać nowych diagramów. Zmiany w strukturze powiązań, a zatem i we wskaźniku, możemy obliczać bez przesuwania jednostek na diagramie, kreśląc diagram dopiero w przypadku wyboru najlepszego podziału. Wszelkie powiązania można bowiem odczytywać bez trudu z diagramu uporządkowanego w pierwszej wersji. Dokonuje się tego w ten sposób, że klasyfikuje wszystkie powiązania danej jednostki z innymi jednostkami do powiązań wewnętrznych i zewnętrznych (zależnie od przyjętej wersji podziału) oraz do silnych i słabych (nieistotnych), dla każdej wersji podziału obliczając wskaźniki powiązań i wybierając wersję najkorzystniejszą o najwyższych wskaźnikach.

Obliczanie — zgodnie z podanym tu wzorem — takiego współczynnika zagęszczenia (koncentracji) powiązań wymaga wielokrotnie mniej czasu niż obliczanie na przykład współczynników zmienności. Jednocześnie uzyskane wyniki są podobne.

Takim wskaźnikiem posługiwałem się z powodzeniem przy rejonizacji produkcyjno-rolniczej indywidualnej gospodarki chłopskiej powiatu lubelskiego. Poniżej podam dla ilustracji wyniki uzyskane dzięki zastosowaniu wskaźnika koncentracji powiązań, oraz dla porównania kształtowanie się współczynników zmienności. Kwestia dotyczyła ewentualnego dołączenia do któregoś z rejonów miasta Świdnika leżącego na granicy obu rejonów i do obu ciążącego.

Rejonizacja przeprowadzona dla lat 1966—1968 (średnie trzyletnie) opierała się na dziewięciu cechach diagnostycznych przeliczonych na 100 ha użytków rolnych. Były to mianowicie: powierzchnia trwałych użytków zielonych (1), pszenicy (2), żyta (3), buraków cukrowych (4), warzyw (5), motylkowych drobnonasiennych (6), przemysłowych poza burakami cukrowymi (7) oraz obsada w sztukach dużych trzodą chlewną (8) i bydłem (9). Różnice obliczano według wzoru:

$$R_{1,2} = \sum_{j=1}^{n} \left| \frac{a_{1j} - a_{2j}}{M_j} \right| \cdot 100$$

Skala, która służyła ostatecznemu porządkowaniu, była czterostopniowa, a mianowicie: I stopień do 270, II od 270 do 360 (te dwa stopnie powiązań uznawałem za bliskie), III od 360 do 540 i IV powyżej 540 (te dwa stopnie to powiązania dalekie). Miasto Świdnik ciążyło do dwu rejonów (można też je traktować jako wydzielające się, gdyż powiązań bliższych rzędów nie ma dla niego wiele):

1 — rejonu Głuska (gromady: Dąbrowica, Głusk, Jastków, Kawęczyn, Konopnica, Kozice Dolne, Krasienin, Piaski, Tomaszowice, Zemborzyce);

2 — rejonu Łęcznej (gromady: Ciecierzyn, Krępiec, Łęczna, miasto Łęczna, Łuszczów, Mełgiew, Wólka).

Na podstawie wykreślonego diagramu Czekanowskiego obliczono wskaźnik zagęszczenia powiązań dla dwóch możliwych rozwiązań: miasto Świdnik w składzie rejonu Głuska i miasto Świdnik w składzie rejonu Łęcznej. Wskaźniki te przedstawia tabela 2.

Kształtowanie się wskaźników zagęszczenia powiązań jest korzystniejsze w drugim wariancie, gdyż spadkowi wskaźnika w rejonie Łęcznej towarzyszy znacznie silniejszy jego wzrost w rejonie Głuska (i to zarówno w wielkościach bezwzględnych, jak i względnych). Współczynniki zmienności dla obu wariantów przedstawia tabela 3.

---

[15] Posługuję się tu co prawda w miejsce stosowanych przez Z. Gługiewicza średnich współczynników zmienności ($\bar{V}_x \frac{\Sigma V_x}{N}$), wielkościami sum współczynników zmienności, tj. ΣV_x. Nie zmienia to istoty samego kryterium. Unika się w ten sposób po prostu zbędnych dodatkowych dzieleń przez liczbę cech (N).

Tab. 2

Wariant	Rejon Głuska			Rejon Łęcznej		
	powiązania w % ogółu możliwych powiązań					
	rząd powiązań	wewn. rejonu	poza rejonem	rząd powiązań	wewn. rejonu	poza rejonem
I — Świdnik w rejonie Głuska	I+II	81,81	37,76	I+II	95,24	46,22
	III+IV	18,19	62,24	III+IV	4,76	53,78
	Wzp=2,17+3,42=5,59			Wzp=2,06+11,30=13,36		
II — Świdnik w rejonie Łęcznej	I+II	95,56	37,86	I+II	82,14	42,19
	III+IV	4,44	62,14	III+IV	17,86	57,81
	Wzp=2,52+14,00=16,52			Wzp=1,95+3,24=5,19		

Tab. 3

Cecha	Powiat	Wariant I		Wariant II	
		rejon Głuska	rejon Łęcznej	rejon Głuska	rejon Łęcznej
Trw. użytki zielone	89,58	76,98	36,91	74,15	46,81
Pszenica	30,53	20,84	12,10	21,58	11,95
Żyto	28,20	20,41	18,12	21,30	18,91
Buraki cukrowe	42,04	26,40	24,49	26,55	23,57
Warzywa	65,77	73,81	50,97	29,59	60,22
Motylkowe	34,68	12,51	16,10	12,60	17,37
Przemysłowe	100,54	54,41	64,75	47,54	62,71
Trzoda chlewna	21,03	11,51	17,50	11,70	13,19
Bydło	14,10	10,28	14,79	9.72	15,43
V_x	426,47	307,15	255,73	254,73	270,16

Widoczne jest zatem, że wskaźnik zagęszczenia powiązań jest nader pomocny dla prawidłowego dokonania podziału. W szczególności ułatwia on zaliczanie jednostek wątpliwych do któregoś z wydzielonych rejonów, czyli umożliwia najbardziej prawidłowe — przy danej przyjętej skali — podzielenie badanej zbiorowości.

Dla uchwycenia, jakie wielkości przybiera wskaźnik koncentracji powiązań, pokusiłem się o obliczenie takich wskaźników dla przeprowadzonych już podziałów, uznanych przy pomocy pewnych kryteriów za dobre. Dotyczy to diagramów z prac A. Fajferka, W. Kwiecienia, J. Liczkowskiego.[16] Wielkości wskaźnika kształtują się najczęściej powy-

[16] Por. F a j f e r e k: *op. cit.;* K w i e c i e ń: *op. cit.;* L i c z k o w s k i: *op. cit.*

żej 10. Dosyć często zdarza się jednak, że w danym podziale dla jednego z wydzielonych rejonów wskaźnik jest niższy od 10, dla pozostałych — wyższy. W żadnym przypadku wielkość jego nie spada poniżej 5. W odniesieniu do tego jednego rejonu ($W_{zp} < 10$) wiąże się to chyba z tym, że na ogół dokonujący rejonizacji niechętnie pozostawiają słabiej łączące się jednostki w ogóle poza rejonami. Przy udziale tych słabiej łączących się jednostek tworzy się z reguły rejon o niższym wskaźniku. Można by było więc przyjąć 10 za graniczną wielkość wskaźnika dla prawidłowego podziału (z tym, że dla jednego rejonu może on ewentualnie wahać się w granicach od 5 do 10). Jeśli wskaźniki zagęszczenia powiązań kształtują się poniżej tej granicy i żadne zmiany w podziale (przesunięcia jednostek) nie powodują korzystniejszego ich ukształtowania, wskazuje to na konieczność dokonania zmiany w skali (granicach przedziałów).

РЕЗЮМЕ

Первая часть статьи посвящена проблематике расчета средней арифметической и стандартного отклонения в таксономическом методе разностей Я. Чекановского, примененного в пространственных исследованиях. Эти величины, по мнению автора, не являются характеристиками для исследуемой совокупности, они служат только самому методу, что вызывает как бы их фикцийность.

Во второй части представляется введенный автором (для улучшения раздела) упрощенный показатель. Автор называет его показателем сгущения связей. Достоинством его является простота расчета и возможность расчета из первого варианта диаграммы (без трудоемкого черчения новых диаграмм).

Автор приводит результаты использования представленного показателя, которые совпадают с результатами, опирающимися на коэффициенты вариантности.

ZUSAMMENFASSUNG

Im ersten Teil des Aufsatzes wird die Problematik der Errechnung des arithmetischen Mittels und der Standardabweichungen bei der taxonomen Differentialmethode von J. Czekanowski erörtert, in der Anwendung auf räumliche Forschungen. Nach dem Verfasser sind diese Werte keine Charakteristiken der untersuchten Gruppe, sie dienen nur der Methode selbst und sind demzufolge gewissermassen fiktiv.

Im zweiten Teil stellt der Verfasser den von sich selber eingeführten — zur Erleichterung der Optimalisierung der Einteilung vereinfachten — Index dar. Dieser wird von ihm Index der eingedichteten Zusammenhänge genannt. Seine Vorteile sind Einfachheit in den Berechnungen und die Möglichkeit, auf Grund des ersten Diagramms zu berechnen (ohne das sehr zeitraubende Entwerfen weiterer Diagramme).

Der Verfasser bringt Ergebnisse der Arbeit mit dem dargestellten Index, die mit den auf Grund von Variabilitätskoeffizienten ermittelten Werten übereinstimmen.

Pap. druk. sat. III kl. 80 g	Format B5 (70×100)	Stron druku: 14
Annales UMCS, Lublin 1972	Drukarnia Uniwersytecka w Lublinie	Zam. nr 241 z dn. 15 VI 72
Nakład: 600+50 egz. — B-4	Maszynopis otrzymano 15 VI 1972 r.	Druk ukończ.: XII 1972 r.